

基于余弦相似度和图卷积网络的电力负荷预测方法

Ji Shan¹, 姜巍², 景鑫¹

(1. 杭州迈拓大数据服务有限公司, 杭州 310007;

2. 国网浙江省电力有限公司科技创新中心, 杭州 310007)

摘要: 针对现有电力负荷预测模型难以深入提取时空关联特征, 模型泛化能力弱, 无法同时胜任短期和长期的电力负荷预测的问题, 提出一种面向多用户的基于余弦相似度和全局-局部协同图卷积网络的电力负荷预测方法。首先, 利用余弦相似度来学习不同节点负荷数据之间的相似模式, 以提取深层次的时空关联特征。其次, 对影响电力负荷变化趋势的静态全局因素和动态局部因素进行协同建模, 以提升模型的泛化能力。最后, 通过在一个实测数据集上进行的大量实验, 验证了该方法在短期和长期负荷序列预测任务中同时具备有效性和稳定性。

关键词: 多节点电力负荷预测; 时空关联特征; 余弦相似度; 图卷积

DOI: 10.19585/j.zjdl.202501007

开放科学(资源服务)标识码(OSID):



A power load forecasting method using cosine similarity and a graph convolutional network

Ji Shan¹, JIANG Wei², JING Xin¹

(1. Hangzhou Data Far East Technology Pty. Ltd., Hangzhou 310007, China;

2. Center of Technology and Innovation State Grid Zhejiang Electric Power Co., Ltd., Hangzhou 310007, China)

Abstract: To address the challenges of existing power load forecasting models, which struggle to deeply extract spatiotemporal correlation features and exhibit weak generalization capabilities—failing to simultaneously manage both short-term and long-term forecasting—this study proposes a multi-user power load forecasting method using cosine similarity and a global-local collaborative graph convolutional network. First, cosine similarity is utilized to learn similar patterns between load data from different nodes, allowing for the deep extraction of spatiotemporal correlation features. Second, a collaborative modeling approach is applied to static global factors and dynamic local factors that influence power load trends, enhancing the model's generalization ability. Finally, extensive experiments on a real-world dataset demonstrate the method's effectiveness and robustness in forecasting both short-term and long-term load series.

Keywords: multi-node power load forecasting; spatiotemporal correlation features; cosine similarity; graph convolutional

0 引言

电力负荷预测是制定电力系统调度决策的重要依据, 它在电力系统实时状态分析、电力系统安全风险评估、配电优化平衡等多个应用场景中起到了至关重要的作用, 为电力系统的安全稳定运行提供重要保障^[1-5]。近年来, 随着分布式电源在电网中的比重逐渐上升, 基于局部区域内的多节点电力负荷预测逐渐成为研究热点。然而, 由

于用户侧用电行为的多样性和用电模式的复杂性, 实现准确的电力负荷预测极具挑战。

深度学习在处理具有复杂模式的高维非线性负荷序列数据方面展现出强大的特征提取能力, 因此在电力负荷预测领域得到越来越广泛的应用。学者们主要基于RNN(循环神经网络)^[6]进行了一系列改进工作。文献[7]基于RNN引入记忆单元和门控机制, 提出了LSTM(长短期记忆网络), 有效缓解了梯度消失问题, 并能有效地学习序列的历史依赖关系。文献[8]提出了BiLSTM(双向长短期记忆网络), 拥有更丰富的信息表示能力和更好的上下文理解能力。文献[9]提出了一种基于

自回归移动平均法和LSTM的混合模型, 结合了线性预测和非线性预测, 能够有效地捕获负荷数据的变化规律。然而, 上述基于RNN的扩展研究主要关注如何提取负荷数据的时序特征, 忽略了多维负荷数据的时空关联特征, 导致预测精确度和准确度等性能不佳。

为了提取多维负荷数据的时空关联特征, 学者们提出了多种新的时空模型, 以在时间和空间维度上同时对多维负荷数据进行建模^[10-11]。CNN(卷积神经网络)^[12]因其在处理欧氏空间结构数据上的优越性, 被广泛用于提取多维负荷数据的空间特征。例如, 文献[13]提出了一种CNN-LSTM模型, 以同时提取负荷数据的空间特征和时序特征。文献[14]提出了一种CNN-BiLSTM模型, 进一步提升了模型的特征提取能力。然而, 此类基于CNN的时空模型将不同节点负荷数据之间的关联性假设为固定模式, 因此难以深入提取多维负荷数据的时空关联特征, 从而限制了预测性能的提升。

GCN(图卷积网络)^[15]能够有效地处理非欧氏空间结构数据, 因此被广泛应用于交通流预测等领域。利用GCN提取交通流数据的时空关联特征的经典方法, 通常基于节点之间地理位置的邻近度来构建邻接矩阵^[16]。然而, 在多节点电力负荷预测任务中, 上述基于节点地理位置构建的邻接矩阵难以有效表达不同节点负荷数据之间的关联性。因为该方法存在着以下两点缺陷: 模型过度依赖节点地理位置的邻近度, 导致难以利用具有相似模式的不同节点负荷数据之间的关联信息; 基于节点地理位置生成的邻接矩阵通常是静态的, 在一定程度上限制了模型捕获时空关联特征的动态变化模式。文献[17]提出了GCNLS-STLF(结合LSTM和自注意力机制的图卷积网络), 利用LSTM将原始多维负荷数据映射为包含不同用户负荷数据之间关联信息的电力负荷图, 并结合GCN来提取时空关联特征, 从而提升电力负荷的预测精度。该研究揭示了利用时空模型提升多节点电力负荷预测准确性的关键在于: 基于多维负荷数据所蕴含的序列信息来学习不同节点负荷数据之间的隐式交互模式, 并将其用于提取时空关联特征。

受此启发, 本文提出了一种新的时空模型, 即CS-GLGCN(基于余弦相似度的全局-局部协同图卷积网络)。相较于GCNLS-STLF, 本文利用了不同节点负荷数据之间的余弦相似度来提取关联信息, 并对影响电力负荷变化趋势的静态全局因素和动态局部因素进行协同建模, 以提取不同时间尺度下的时空关联特征。实验证明, 本文所提出的方法有效提升了模型的准确度并增强了模型的泛化能力。本文主要贡献如下:

1) 提出一种基于余弦相似度的图学习方法, 能够基于多维负荷数据所蕴含的序列信息, 利用余弦相似度来学习不同节点负荷数据之间的相似模式, 为不同节点负荷数据之间的关联性赋予权重, 进而生成包含负荷关联信息的负荷相似图并将其作用于空间特征的聚合, 从而提取到深层次的时空关联特征。

2) 提出一种全局-局部协同的图卷积方法, 利用多维负荷数据在不同时间尺度下的负荷相似图, 结合图卷积网络对影响电力负荷变化趋势的静态全局因素和动态局部因素进行协同建模, 以提取不同时间尺度下的时空关联特征, 进而提升模型在面对不同数据分布和预测需求时的泛化能力。

1 算法原理

1.1 问题定义

本文研究的多节点电力负荷预测任务可以描述为: 学习一种非线性映射函数 f , 基于负荷相似图 G 和时间步长为 T 的历史负荷特征矩阵 $X \in R_{N \times T}$, 预测时间步长为 T 的未来负荷状态 $Y \in R_{N \times T}$, 其中 N 代表节点数量。预测过程如式(1)所示:

$$[X_{t-T+1}, X_{t-T+2}, \dots, X_t; G] \xrightarrow{f} [Y_{t+1}, Y_{t+2}, \dots, Y_{t+T}] \quad (1)$$

式中: $X_t \in R_{N \times 1}$ 为 X 在历史某个时间步 t 时的负荷特征矩阵; $Y_{t+T} \in R_{N \times 1}$ 为 Y 在未来某个时间步 $(t+T)$ 时的负荷状态; T 为输入和预测的时间步长。

1.2 余弦相似度

余弦相似度是一种能够在高维特征空间中衡量两个向量相似度的方法^[18]。时间步长均为 T 的负荷序列 L_i 和 L_j 之间的余弦相似度 $S_{i,j,T}$ 的计算方

法,如式(2)所示:

$$S_{i,j,T} = \frac{\sum_{l=1}^T (L_{i,l} L_{j,l})}{\sqrt{\sum_{l=1}^T (L_{i,l})^2} \cdot \sqrt{\sum_{l=1}^T (L_{j,l})^2}} \quad (2)$$

式中: $L_{i,l}$ 和 $L_{j,l}$ 分别为序列 L_i 和 L_j 中的第 l 个元素, $i, j \in (1, 2, 3, \dots, N)$ 。

基于节点数为 N 、时间步长为 T 的负荷特征矩阵 $X \in R_{N \times T}$, 本文利用余弦相似矩阵计算函数 $CosineSimilarity(\bullet)$ 生成负荷相似图 $G \in R_{N \times N}$, 如式(3)所示:

$$G = CosineSimilarity(X) = \begin{bmatrix} S_{1,1,T}, S_{1,2,T}, \dots, S_{1,N,T} \\ S_{2,1,T}, S_{2,2,T}, \dots, S_{2,N,T} \\ \vdots \\ S_{N,1,T}, S_{N,2,T}, \dots, S_{N,N,T} \end{bmatrix} \quad (3)$$

1.3 图卷积网络

GCN 通过在所有节点上共享权重以学习图中不同位置节点间的相似模式。经典的 GCN 运算模式^[15]如式(4)所示:

$$H = \tilde{D}^{-\frac{1}{2}} \tilde{A} \tilde{D}^{-\frac{1}{2}} XW \quad (4)$$

式中: X 和 H 分别为输入和输出的特征矩阵; $\tilde{A}$ 为邻接矩阵 A 加上单位矩阵 I (考虑到自连接); $\tilde{D}$ 为 $\tilde{A}$ 的度矩阵; W 为权重矩阵。

2 基于余弦相似度的全局-局部协同图卷积网络

2.1 总体架构

CS-GLGCN 模型如图 1 所示, 主要由以下 5 个模块构成。

1) 输入模块将训练数据整体输入给基于全局余弦相似度的图学习组件, 同时将分批处理后的数据批量输入给线性层组件和基于局部余弦相似度的图学习组件。该模块包括 1 个整体输入层组件和 1 个批量输入层组件。

2) 时间门控卷积模块提取批量输入数据的时序特征, 包括 2 个 TCL (时间卷积层) 组件。

3) 图学习模块基于不同时间尺度的多维负荷数据生成 S-LSG (静态负荷相似图) 和 D-LSG (动态负荷相似图), 分别用于表征负荷之间的全局与局部相似权重。该模块包括 1 个基于全局余弦相似度的图学习组件和 1 个基于局部余弦相似度的图学

习组件。

4) 图卷积模块将时序特征分别与 S-LSG 和 D-LSG 相结合, 以提取多维负荷数据的全局和局部的时空关联特征。该模块包括 1 个用于提取全局时空关联特征的图卷积组件和 1 个用于提取局部时空关联特征的图卷积组件。

5) 输出模块对初步提取的粗粒度时序特征、全局时空关联特征和局部时空关联特征进行编码, 得到相应的细粒度特征, 并通过特征聚合来得到细粒度的时空融合特征, 再利用多层感知机对细粒度的时空融合特征进行解码, 从而得到最终的预测结果。该模块包括 1 个时间编码器组件、2 个空间编码器组件、1 个特征聚合器组件和 1 个多层感知机解码器组件。

2.2 时间门控卷积模块

本文在 TCL-A 组件和 TCL-B 组件中利用 CNN 来处理输入的多维负荷数据。扩张因果卷积^[19]能够在保持卷积核尺寸不变的情况下, 通过扩大卷积核元素之间的间距以增大感受野, 从而捕获更大尺度的时间依赖关系。此外, 扩张因果卷积采用非递归的方式来处理负荷序列, 可有效缓解模型在处理长期负荷序列时出现的梯度爆炸问题。因此, 本文在 CNN 的基础上融合了扩张因果卷积, 以提升模型对长期负荷序列的处理能力。

首先, 时间门控卷积模块通过扩张因果卷积将批量输入的第 t 时间步、节点数量为 N 、时间步长为 T 的局部负荷特征矩阵 $X_{t-t+T-1}^{local} \in R_{N \times T}$ 映射为 $X_{dc1} \in R_{N \times (T-K_{dc}+1)}$ 和 $X_{dc2} \in R_{N \times (T-K_{dc}+1)}$, 其中 K_{dc} 代表因果卷积核的大小。然后, 该模块利用一种基于双激活函数的门控机制^[20]来处理映射后的特征。该机制能够有效控制时间卷积层中的序列信息流^[21], 从而提取出多维负荷数据的局部时序特征 $H_{time} \in R_{N \times (T-K_{dc}+1)}$, 如式(5)所示:

$$H_{time} = \sigma(X_{dc1} \times \theta_1 + b_1) \odot g(X_{dc2} \times \theta_2 + b_2) + e \quad (5)$$

式中: $\sigma(\bullet)$ 为 sigmoid 激活函数, 决定了信息传递的比例; $g(\bullet)$ 为 tanh 激活函数, 用于处理输出的隐藏特征; $\times$ 为矩阵乘积, $\odot$ 为哈达玛乘积; θ_1 、 θ_2 、 b_1 、 b_2 为模型的可学习参数; e 为正态分布的高斯白噪声, 用于降低时序特征的过拟合风险。

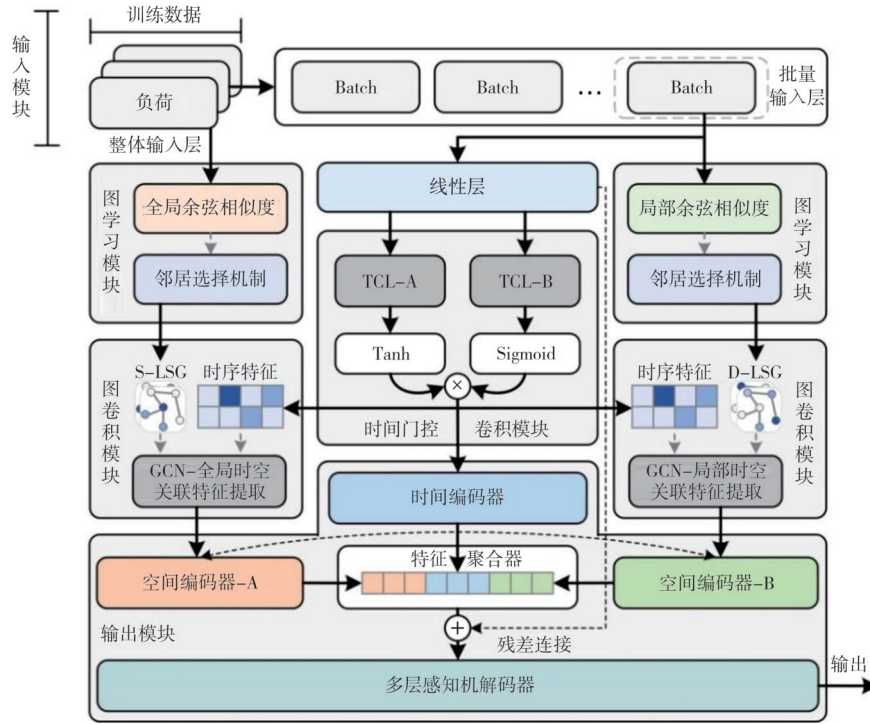


图 1 CS-GLGCN 模型的总体架构

Fig.1 General architecture of the CS-GLGCN model

2.3 图学习模块

图学习模块基于多维负荷数据所蕴含的序列信息, 利用余弦相似度来学习不同节点负荷数据之间的全局相似模式和局部相似模式, 以分别表征不同节点负荷数据之间的全局相似权重与局部相似权重。如图 1 所示, 基于全局余弦相似度的图学习组件利用节点数量为 N 、时间步长为 T_s 的全局负荷特征矩阵 $X_{1 \sim T_s}^{\text{global}} \in R_{N \times T_s}$ 来生成 S-LSG (即 $G_{\text{global}} \in R_{N \times N}$), 并对其进行邻居选择从而得到 $\hat{G}_{\text{global}} \in R_{N \times N}$, 如式(6)和式(7)所示:

$$G_{\text{global}} = \text{CosineSimilarity}(X_{1 \sim T_s}^{\text{global}}) \quad (6)$$

$$\hat{G}_{\text{global}} = \text{Selector}[k_g](G_{\text{global}}) \quad (7)$$

式中: T_s 为全局时间步长; $\text{Selector}[k_g](\bullet)$ 为全局邻居选择函数; k_g 为选择的邻居数, 即以每个节点为中心, 保留与其全局负荷模式最相似的前 k_g 个权重, 并将其余 $(N - k_g)$ 个权重值置零。

同时, 在模型训练过程中, 基于局部余弦相似度的图学习组件利用局部负荷特征矩阵 $X_{t \sim t+T-1}^{\text{local}}$ 来生成 D-LSG (即 $G_{\text{local}} \in R_{N \times N}$), 并对其进行邻居选择从而得到 $\hat{G}_{\text{local}} \in R_{N \times N}$, 如式(8)和式(9)所示:

$$G_{\text{local}} = \text{CosineSimilarity}(X_{t \sim t+T-1}^{\text{local}}) \quad (8)$$

$$\hat{G}_{\text{local}} = \text{Selector}[k_l](G_{\text{local}}) \quad (9)$$

式中: T 为局部时间步长, $\text{Selector}[k_l](\bullet)$ 为局部邻居选择函数; k_l 为选择的邻居数, 即以每个节点为中心, 保留与其局部负荷模式最相似的前 k_l 个权重, 其余 $(N - k_l)$ 个权重值置零。

2.4 图卷积模块

图卷积模块对影响电力负荷变化趋势的静态全局因素和动态局部因素进行协同建模, 从而提取出多维负荷数据在不同时间尺度下的时空关联特征, 以获取更加泛化的特征表示。如图 1 所示, 该模块将时间门控卷积模块提取的时序特征 H_{time} 分别与 $\hat{G}_{\text{global}}$ 和 $\hat{G}_{\text{local}}$ 进行图卷积运算, 以提取全局时空关联特征 $H_{\text{global}} \in R_{N \times (T - K_{\text{dc}} + 1)}$ 和局部时空关联特征 $H_{\text{local}} \in R_{N \times (T - K_{\text{dc}} + 1)}$, 如式(10)和式(11)所示:

$$H_{\text{global}} = \hat{G}_{\text{global}} H_{\text{time}} W_0 \quad (10)$$

$$H_{\text{local}} = \hat{G}_{\text{local}} H_{\text{time}} W_1 \quad (11)$$

式中: W_0 和 W_1 为可学习权重矩阵。

2.5 输出模块

如图 1 所示, 输出模块采用时间编码器组件、空间编码器-A 组件和空间编码器-B 组件分别对提取的粗粒度时序特征 H_{time} 、全局时空关联特征

H_{global} 和局部时空关联特征 H_{local} 进行编码, 得到相应的细粒度特征 $\hat{H}_{\text{time}} \in \mathbb{R}^{N \times E_{\text{EncodeSize}} \times C_{\text{out}}}$ 、 $\hat{H}_{\text{global}} \in \mathbb{R}^{N \times E_{\text{EncodeSize}} \times C_{\text{out}}}$ 和 $\hat{H}_{\text{local}} \in \mathbb{R}^{N \times E_{\text{EncodeSize}} \times C_{\text{out}}}$, 如式(12)–(14)所示:

$$\hat{H}_{\text{time}} = \text{TimeEncoder}(H_{\text{time}}) \quad (12)$$

$$\hat{H}_{\text{global}} = \text{SpatialEncoderA}(H_{\text{global}}) \quad (13)$$

$$\hat{H}_{\text{local}} = \text{SpatialEncoderB}(H_{\text{local}}) \quad (14)$$

式中: C_{out} 为输出通道大小; $\text{TimeEncoder}(\cdot)$ 和 $\text{SpatialEncoderB}(\cdot)$ 为卷积核大小为1的单层CNN; $\text{SpatialEncoderA}(\cdot)$ 为卷积核大小为3的单层CNN; $E_{\text{EncodeSize}}$ 为编码大小。

然后, 特征聚合器组件对提取的细粒度特征进行聚合, 得到细粒度时空融合特征 $Z \in \mathbb{R}^{N \times (E_{\text{EncodeSize}} \times 3) \times C_{\text{out}}}$, 如式(15)所示:

$$Z = \text{Aggregation}(\hat{H}_{\text{time}}, \hat{H}_{\text{global}} W_{\text{adp}}, \hat{H}_{\text{local}} (I - W_{\text{adp}})) \quad (15)$$

式中: $\text{Aggregation}(\cdot)$ 为特征聚合函数; I 为单位矩阵; W_{adp} 为自适应权重。

通过自适应地调整全局和局部时空关联特征的权重, 模型能够更灵活地适应不同的数据分布和预测需求。

考虑到多层感知机能够有效地映射高维特征空间中序列之间深层的依赖关系^[22], 因此输出模块的最后部分采用了多层感知机对细粒度时空融合特征 Z 进行解码并输出最终的预测结果 $\hat{Y} \in \mathbb{R}^{N \times T \times C_{\text{out}}}$, 如式(16)所示:

$$\hat{Y} = \text{MLP}(Z + r_{\text{residual}}) \quad (16)$$

式中: $\text{MLP}(\cdot)$ 为多层感知机解码器组件; r_{residual} 为 X_{t-T+1}^{local} 的线性残差。

3 算例分析

采用中国某地区电力公司收集的实测多维负荷数据(节点个数为35, 采样周期为15 min, 采样点为8 832个)进行仿真实验验证。采用Z-Score归一化方法来处理输入数据, 并将数据集分为3部分: 70%用于训练, 10%用于验证, 20%用于测试。

本文所提出CS-GLGCN模型在Pytorch环境下使用Python语言实现, 所有实验都在一台处理器为13th Gen Intel(R) Core(TM)i7-13700KF@

3.40 GHz、显卡为NVIDIA GeForce RTX 4060Ti GPU的计算服务器上运行。时间门控卷积层的输出通道被设置为32, 因果卷积的内核大小 K_{dc} 被设置为3, 随机初始化节点嵌入的维度设置为32。图学习模块中的邻居数 k_g 和 k_l 被设置为6。图卷积模块的输出通道被设置为16, 图卷积层数被设置为2。输出模块中的编码大小 EncodeSize 被设置为256。同时采用Adam优化器训练所提出的模型, 初始学习率被设置为0.000 3, 衰减率被设置为0.000 5。批量大小被设置为32, 训练代数被设置为100, 并应用了早停法。

3.1 基线与基准

本节将CS-GLGCN模型与LSTM^[7]、CNN-LSTM^[13]、DLinear^[23]、PatchTST^[24]和GCNLS-STLF^[17]5种电力负荷预测模型进行比较, 以验证CS-GLGCN模型在短期负荷序列预测和长期负荷序列预测任务中同时具有有效性和稳健性。

采用 r_{mae} (平均绝对误差)、 r_{mape} (平均绝对百分比误差) 和 r_{rmse} (均方根误差) 来评估本文提出的CS-GLGCN模型和上述5种对比模型的预测性能。

$$r_{\text{mae}} = \frac{1}{n} \sum_{i=1}^n |Y_i - \hat{Y}_i| \quad (17)$$

$$r_{\text{mape}} = \frac{1}{n} \sum_{i=1}^n \left| \frac{Y_i - \hat{Y}_i}{Y_i} \right| \times 100\% \quad (18)$$

$$r_{\text{rmse}} = \sqrt{\frac{1}{n} \sum_{i=1}^n (Y_i - \hat{Y}_i)^2} \quad (19)$$

式中: n 、 Y_i 、 $\hat{Y}_i$ 分别为测试集的样本量、真实值和预测值。

3.2 预测性能比较

本节比较6种模型在预测时间步长(1个步长为15 min)为3、6、12(@3、@6、@12)时的短期负荷序列预测性能; 预测时间步长为96、192、288(@96、@192、@288)时的长期负荷序列预测性能。表1、表2记录了6种模型3次预测实验的平均结果。

从表1、表2中的实测数据得出以下结论:

1)CS-GLGCN模型相比于次优模型, 在预测步长为3时, r_{mae} 和 r_{mape} 分别提升了2.630%、2.087%; 在预测步长为6时, r_{mae} 、 r_{mape} 、 r_{rmse} 分别提升了2.675%、2.815%、0.519%。结果表明,

表 1 6 种模型的短期负荷序列预测性能比较

Table 1 Comparison of short-term load series forecasting performance of six models

模型	@3			@6			@12		
	r_{mae}	$r_{\text{mape}}/\%$	r_{rmse}	r_{mae}	$r_{\text{mape}}/\%$	r_{rmse}	r_{mae}	$r_{\text{mape}}/\%$	r_{rmse}
LSTM	37.21	12.00	53.49	48.01	15.71	69.79	50.19	19.07	76.37
CNN-LSTM	34.25	11.21	<u>49.92</u>	36.95	12.36	54.58	40.83	14.74	61.70
DLinear	48.81	16.30	78.48	71.42	26.49	108.96	63.23	26.52	97.486
PatchTST	48.19	16.80	75.06	71.15	27.94	101.94	73.44	28.76	104.28
GCNLS-STLF	<u>33.84</u>	<u>11.02</u>	48.70	<u>35.51</u>	<u>12.08</u>	<u>53.97</u>	<u>35.90</u>	12.39	56.58
CS-GLGCN	32.95	10.79	52.98	34.56	11.74	53.69	35.75	<u>12.95</u>	<u>56.75</u>

注：每列的最优值以粗体标记，次优值以下划线标记。

表 2 6 种模型的长期负荷序列预测性能比较

Table 2 Comparison of long-term load series forecasting performance of six models

模型	@96			@192			@288		
	r_{mae}	$r_{\text{mape}}/\%$	r_{rmse}	r_{mae}	$r_{\text{mape}}/\%$	r_{rmse}	r_{mae}	$r_{\text{mape}}/\%$	r_{rmse}
LSTM	52.29	17.62	80.79	56.73	18.98	86.12	57.55	19.42	87.89
CNN-LSTM	46.73	16.82	73.20	61.37	18.62	95.54	53.64	17.29	82.39
DLinear	<u>34.91</u>	<u>15.27</u>	64.72	31.93	12.11	56.09	33.32	11.84	60.06
PatchTST	35.23	16.16	<u>63.33</u>	<u>31.61</u>	<u>11.70</u>	<u>54.74</u>	<u>32.95</u>	<u>11.59</u>	<u>58.39</u>
GCNLS-STLF	42.63	16.47	69.90	46.79	16.91	74.36	49.81	17.15	79.63
CS-GLGCN	31.40	11.34	54.78	30.49	10.54	53.56	32.54	10.88	58.27

注：每列的最优值以粗体标记，次优值以下划线标记。

CS-GLGCN模型具有更优秀的短期负荷序列预测性能。

2)CS-GLGCN模型相比于次优模型，在预测步长为 96 时， r_{mae} 、 r_{mape} 、 r_{rmse} 分别提升了 10.054%、25.737%、13.501%；在预测步长为 192 时， r_{mae} 、 r_{mape} 、 r_{rmse} 分别提升了 3.543%、9.915%、2.156%；在预测步长为 288 时， r_{mae} 、 r_{mape} 、 r_{rmse} 分别提升了 1.244%、6.126%、0.206%。结果表明，CS-GLGCN模型的长期负荷序列预测性能均优于 5 种对比模型。

综上，CS-GLGCN模型在短期和长期负荷序列预测任务中均具有先进的预测性能。

3.3 消融实验

通过比较CS-GLGCN模型的 3 种变体，以衡量不同组件对CS-GLGCN模型预测性能的贡献，如下所示：

1)CS-N(w/oglc)：去除全局、局部图学习模块和全局-局部协同图卷积模块，即仅保留时序特征聚合的CS-GLGCN模型。

2)CS-LN(w/og)：去除全局图学习模块和全局图卷积模块，即仅保留时序特征和局部时空关

联特征聚合的CS-GLGCN模型。

3)CS-GN(w/ol)：去除局部图学习模块和局部图卷积模块，即仅保留时序特征和全局时空关联特征聚合的CS-GLGCN模型。

为清晰地展示不同组件对CS-GLGCN模型的贡献，分别在预测时间步长为 12 和 96 时进行了消融实验。表 3 记录了每种变体 3 次实验的平均结果。

表 3 CS-GLGCN模型 3 种变体的预测性能比较

Table 3 Comparison of the forecasting performance of three variants of the CS-GLGCN model

模型	@12			@96		
	r_{mae}	$r_{\text{mape}}/\%$	r_{rmse}	r_{mae}	$r_{\text{mape}}/\%$	r_{rmse}
CS-N(w/oglc)	46.09	17.21	72.14	39.82	16.79	68.23
CS-LN(w/og)	42.37	16.08	66.98	<u>34.66</u>	<u>13.99</u>	<u>60.93</u>
CS-GN(w/ol)	<u>40.66</u>	<u>15.06</u>	<u>65.28</u>	34.75	14.09	61.24
CS-GLGCN	35.75	13.02	56.89	32.46	11.65	57.21

注：每列的最优值以粗体标记，次优值以下划线标记。

从表 3 中可得以下结论：

1)在预测步长为 12 和 96 时，CS-GLGCN 模型的 r_{mae} 、 r_{mape} 、 r_{rmse} 均优于CS-N(w/oglc)，表明

基于多维负荷数据所蕴含的序列信息,利用余弦相似度来学习不同节点负荷数据之间的相似模式,并将其作用于空间特征的聚合,能够提取到深层次的时空关联特征,进而提升模型的预测性能。

2)在预测步长为12和96时,CS-LN(w/og)和CS-GN(w/ol)的表现相比于CS-N(w/oglc)有所提升,但整体表现劣于CS-GLGCN模型,表明对影响电力负荷变化趋势的静态全局因素和动态局部因素进行协同建模,可有效提升模型在面对不同数据分布和预测需求时的泛化能力。

3)CS-GN(w/ol)相比于CS-LN(w/og),在时间步长为12时, r_{mae} 、 r_{mape} 、 r_{rmse} 分别提升了4.036%、6.343%、2.538%,表明在对周期性不完整的短期负荷序列进行预测时,静态全局时空关联特征相较于动态局部时空关联特征,其提供的稳定关联信息能够更显著地增强模型短期负荷序列预测的稳健性。

4)CS-LN(w/og)相比于CS-GN(w/ol),在时间步长为96时, r_{mae} 、 r_{mape} 、 r_{rmse} 分别提升了0.259%、0.710%、0.506%,表明在预测具有完整周期性的长期负荷序列时,动态局部时空关联特征相较于静态全局时空关联特征,其提供的实时关联信息能够更显著地提升模型长期负荷序列预测的准确性。

4 结语

本文提出面向多节点电力负荷预测的CS-GLGCN时空模型。首先,针对多维负荷数据蕴含丰富序列信息的特点,提出了基于余弦相似度的图学习方法,该方法能够学习不同节点负荷数据之间的相似模式并生成包含负荷关联信息的权重图。然后,将其作用于空间特征的聚合,以提取出深层次的时空关联特征。其次,考虑到静态全局因素和动态局部因素对负荷变化趋势的影响,提出一种全局-局部协同的图卷积方法对上述两类因素进行协同建模,以提升模型在面对不同数据分布和预测需求时的泛化能力。最后,通过在真实的负荷数据集上进行的大量实验,验证了所提出的模型在短期和长期负荷序列预测任务中同时具备有效性和稳定性。

CS-GLGCN模型适用于同一功能区域内多个

不同节点的电力负荷预测,例如住宅区中的不同家庭、商业区中的不同商业店铺、工业区中的不同车间等。这些节点处于同一功能区域内,用电特征具有全局相似性,但由于个体差异性的存在而导致在不同时间段的用电特征具有局部相似性。本文所提出的方法能够有效利用不同节点之间的相似负荷模式来提取时空关联特征,并结合时空图卷积网络进行多节点电力负荷预测。

除多节点电力负荷预测场景,本文所提出的CS-GLGCN模型还可拓展到其他时空预测场景,如多节点燃气消耗预测和多地降水量预测等。未来研究将进一步探索基于余弦相似度的多尺度图学习方法,根据不同的数据分布和预测需求,自适应地调整负荷相似模式的学习范围,以提高多节点电力负荷预测的可靠性。

参考文献

- [1] 沈建良,来骏,张翼,等.基于ACMD和BiGRU-Attention的负荷预测模型研究[J].浙江电力,2023,42(6):70-77.
SHEN Jianliang, LAI Jun, ZHANG Yi, et al. Research on a load forecasting model based on ACMD and BiGRU-Attention[J]. ZheJiang Electric Power, 2023(6):70-77.
- [2] 封钰,宋佑斌,金晟,等.基于随机森林算法和粗糙集理论的改进型深度学习短期负荷预测模型[J].发电技术,2023,44(6):889-895.
FENG Yu, SONG Youbin, JIN Sheng, et al. Improved deep learning model for forecasting short-term load based on random forest algorithm and rough set theory [J]. Power Generation Technology, 2023,44(6):889-895.
- [3] 王涛,朱海南,李丰硕,等.基于高速公路神经网络和注意力机制的短期负荷预测方法[J].电力电容器与无功补偿,2023,44(4):72-79.
WANG Tao, ZHU Hainan, LI Fengshuo, et al. Short-term load forecast method based on highway neural network and attention mechanism [J]. Power Capacitor & Reactive Power Compensation, 2023,44(4):72-79.
- [4] 陈逸枫,张大海,李宇欣,等.基于DWT-MOSMA-SVM的多目标优化短期母线负荷预测[J].电力建设,2023,44(3):49-55.
CHEN Yicong, ZHANG Dahai, LI Yuxin, et al. Multi-objective optimization based on DWT-MOSMA-SVM for short-term bus load forecasting [J]. Electric Power Construction, 2023,44(3):49-55.
- [5] 李振坤,黄滢,李凉,等.计及需求侧响应的主动配电网多时间尺度优化调度[J].电力建设,2023,44(3):36-48.
LI Zhenkun, HUANG Ying, LI Liang, et al. Multi-time

- scale optimal dispatching of active distribution network considering demand-side response[J]. *Electric Power Construction*, 2023, 44(3): 36-48.
- [6] ZAREMBA W, SUTSKEVER I, VINYALS O. Recurrent neural network regularization [J]. *ArXiv e-Prints*, 2014: arXiv:1409.2329.
- [7] MA X L, TAO Z M, WANG Y H, et al. Long short-term memory neural network for traffic speed prediction using remote microwave sensor data [J]. *Transportation Research Part C-emerging Technologies*, 2015, 54: 187-197.
- [8] HUANG Z H, XU W, YU K. Bidirectional LSTM-CRF models for sequence tagging [EB/OL]. [2024-06-02]. <https://arxiv.org/abs/1508.01991v1>.
- [9] 王鑫, 李安桂, 李扬, 等. 基 ARIMA-LSTM 模型的综合能源系统负荷与风光资源预测[J]. *西安建筑科技大学学报(自然科学版)*, 2022, 54(5): 762-769.
- WANG Xin, LI Angui, LI Yang, et al. Multivariate load prediction and wind-solar resource characteristic quantity prediction of integrated energy system based on ARIMA-LSTM model [J]. *Xi'an Univ. of Arch. & Tech. (Natural Science Edition)*, 2022, 54(5): 762-769.
- [10] 黄冬梅, 张宁宁, 胡安铎, 等. 基于双层 XGBoost 和数据增强的空间负荷预测方法[J]. *电力工程技术*, 2023, 42(1): 201-208.
- HUANG Dongmei, ZHANG Ningning, HU Anduo, et al. Spatial load forecasting method based on double-layer XGBoost and data enhancement [J]. *Electric Power Engineering Technology*, 2023, 42(1): 201-208.
- [11] 唐晓, 陈芳, 许强, 等. 改进鲸鱼算法优化的多维度深度极限学习机短期负荷预测[J]. *山东电力技术*, 2023, 50(1): 1-7.
- TANG Xiao, CHEN Fang, XU Qiang, et al. Short-term load forecasting based on multi-dimensional deep extreme learning machine optimized by improved whale algorithm [J]. *Shandong Electric Power*, 2023, 50(1): 1-7.
- [12] LECUN Y, BOTTOU L, BENGIO Y, et al. Gradient-based learning applied to document recognition [J]. *Proceedings of the IEEE*, 86(1998): 2278-2324.
- [13] 陆继翔, 张琪培, 杨志宏, 等. 基于 CNN-LSTM 混合神经网络模型的短期负荷预测方法[J]. *电力系统自动化*, 2019, 43(8): 131-137.
- LU Jixiang, ZHANG Qipei, YANG Zhihong, et al. Short-term load forecasting method based on CNN-LSTM hybrid neural network model [J]. *Automation of Electric Power Systems*, 2019, 43(8): 131-137.
- [14] 王裕鑫. 基于 CNN-BiLSTM 算法的电力负荷预测方法研究[D]. 西安: 西安理工大学, 2022.
- WANG Yuxin. Research on the power load forecasting method based on CNN-BiLSTM [D]. Xi'an: Xi'an University of Technology, 2022.
- [15] KIPF T N, WELING M. Semi-supervised classification with graph convolutional networks [J]. *ArXiv e-Prints*, 2016: arXiv:1609.02907.
- [16] ZHAO L, SONG Y J, ZHANG C, et al. T-GCN: A temporal graph convolutional network for traffic prediction [J]. *IEEE Transactions on Intelligent Transportation Systems*, 2020, 21(9): 3848-3858.
- [17] 史含笑, 王雷春. 结合 LSTM 和自注意力机制的图卷积网络短期电力负荷预测[J]. *计算机应用*, 2024, 44(1): 311-317.
- SHI Hanxiao, WANG Leichun. Short-term power load forecasting by graph convolutional network combining LSTM and self-attention mechanism [J]. *Journal of Computer Applications*, 2024, 44(1): 311-317.
- [18] LUO C J, ZHAN J F, XUE X H, et al. Cosine normalization: using cosine similarity instead of dot product in neural networks [M]// *Lecture Notes in Computer Science*. Cham: Springer International Publishing, 2018: 382-391.
- [19] YU F, KOLTUN V. Multi-scale context aggregation by dilated convolutions [EB/OL]. [2024-06-25]. <https://arxiv.org/abs/1511.07122v3>.
- [20] WU Z H, PAN S R, LONG G D, et al. Graph WaveNet for deep spatial-temporal graph modeling [J]. *ArXiv e-Prints*, 2019: arXiv:1906.00121.
- [21] DAUPHIN N Y, FAN A, AULI M, et al. Language modeling with gated convolutional networks [J]. *ArXiv e-Prints*, 2016: arXiv:1612.08083.
- [22] SHAO Z Z, ZHANG Z, WANG F, et al. Spatial-temporal identity: a simple yet effective baseline for multivariate time series forecasting [C]// *Proceedings of the 31st ACM International Conference on Information & Knowledge Management*. Atlanta GA USA: ACM, 2022: 4454-4458.
- [23] ZENG A L, CHEN M X, ZHANG L, et al. Are transformers effective for time series forecasting [J]. *Proceedings of the AAAI Conference on Artificial Intelligence*, 2023, 37(9): 11121-11128.
- [24] NIE Y Q, NGUYEN N H, SINTHONG P, et al. A time series is worth 64 words: long-term forecasting with transformers [EB/OL]. [2024-06-21]. <https://arxiv.org/pdf/2211.14730>.

收稿日期: 2024-07-03; 修回日期: 2024-09-03

作者简介:

JI Shan(1980), 男, 博士, 主要从事电力系统大数据及人工智能方面的研究工作。

(本文编辑: 张瑞敏)